



Epigenetic inheritance speeds up evolution of artificial organisms

Yoram Vadée-Le-Brun, Jonathan Rouzaud-Cornabas, Guillaume Beslon

► To cite this version:

Yoram Vadée-Le-Brun, Jonathan Rouzaud-Cornabas, Guillaume Beslon. Epigenetic inheritance speeds up evolution of artificial organisms. European Conference on Artificial Life, Jul 2015, York, United Kingdom. 10.7551/978-0-262-33027-5-ch078 . hal-01248383

HAL Id: hal-01248383

<https://inria.hal.science/hal-01248383>

Submitted on 6 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Epigenetic inheritance speeds up evolution of artificial organisms

Yoram Vadée-le-Brun and Jonathan Rouzaud-Cornabas and Guillaume Beslon

Universite de Lyon, CNRS, INRIA Beagle team
INSA-Lyon, LIRIS, UMR5205, F-69621, France
guillaume.beslon@inria.fr

Abstract

DNA is not the sole medium by which parents transmit information to their offspring. Epigenetic inheritance, in particular, is based on the partial transmission of the cellular state of the parental cell to its descendants. Although the reality of epigenetic inheritance is now firmly established, whether it has an influence on the long term evolutionary process is still subject to debate.

To address this question, we used RAevol, an *in silico* experimental evolution platform, and defined 4 scenarios with static or dynamic environments and with or without epigenetic inheritance. Simulations in dynamic environments show that protein inheritance indeed increases the rate of evolution on the long term. But they also show that it impedes evolution in its very first stages. This negative effect can be explained by instabilities generated by the interference between the two inheritance mediums. On the opposite, the long term gain can be explained by protein inheritance reducing the constraints on the genetic regulation network.

Introduction

In biology, inheritance (or heredity) typically refers to the fact that offspring look more like their parents than like other random individuals. We call this kind of inheritance “phenotypic inheritance”. Mendel, in 1866, established the first rules about phenotypic inheritance. In particular he established that some phenotypic traits are encapsulated in a physical “thing” he called allele and that can be transferred from a parent to its offspring. Then, the discovery of DNA structure gave more support to this idea and we began to call these things “genes”. This leads to a subtle semantic switch, *genetic* inheritance progressively becoming the central dogma to explain any kind of phenotypic inheritance.

Although more and more evidence has accumulated against this dogma (Danchin et al., 2011), the notion of non-genetic inheritance (*i.e.*, any kind of phenotypic inheritance which is not due to genetic inheritance) was considered as playing a minor role in the evolutionary dynamics of species until the work of Jablonka et al. (1995). Since then, the precise role of non-genetic inheritance in the evolutionary dynamics of species has been discussed extensively, in particular in the context of “epigenetic inheritance” which is a sub-

set of non-genetic inheritance implying the physical transfer of molecular states from the parents to their offspring.

On the one hand, epigenetic inheritance could increase the fitness of the organisms that are able to transmit some molecular structures to their descendance. The main argument that backs this point of view is the fact that epigenetic inheritance allows for the transmission of an acquired phenotype. It can thus increase the mean fitness in case of rapid environmental change by eliminating the lag-time of the plastic response (Jablonka et al., 1995; Lachmann and Jablonka, 1996; Bonduriansky and Day, 2009). Moreover, as epigenetic inheritance provides heritable phenotypic variation, it creates the possibility to exploit “Stochastic Gene Expression” to survive and maybe adapt to a hostile environmental variation. Indeed, epigenetic inheritance can help transmitting random phenotypic variations that may be less sensitive to this specific environment and thus promote the evolution of new niche-exploitation strategies (Bonduriansky and Day, 2009). Such a behavior has been observed by Adam et al. (2008): some *E. coli* strains become randomly resistant to an antibiotic and can survive and reproduce in this new environment by partially transmitting this resistance trait to their offspring. Finally, epigenetic inheritance allows to increase phenotypic variation by separating the effect of selection on the genotypes and the phenotypes. Indeed, increasing genetic variability could be dangerous because mutations are mostly deleterious and mutations are not reversible. Epigenetic inheritance enables transmission of the adaptive variations while keeping the mutational burden affordable. This can lead to a new strategy for species: when maladapted to an environment they can increase the stochasticity of their gene expression in order to increase their phenotypic variations. Doing so, they increase the probability of finding an adapted phenotype while still being able to transmit it vertically. This could then facilitate evolution toward a new optimum (Pál, 1998; Pál and Hurst, 2004).

This effect could be enhanced by the ontogenesis in multicellular organisms but this is out of scope of our work.

On the other hand, epigenetic inheritance could decrease the rate of evolution. There are two main points supporting

this idea: (i.) If the parent environment and the offspring environment are poorly correlated, non-genetic inheritance can be maladaptive. Indeed, the adaptive plastic response of the parent would interfere with the adaptive plastic response of the offspring in the new environment (Bonduriansky and Day, 2009). (ii.) Non-genetic inheritance is often transmitted with a probability k different of 0.5. This can alter the fair mendelian game and lead to either a maladaptive variation invading the population because of $k > 0.5$ or an adaptive variation lost because of $k < 0.5$ (Pál and Hurst, 2004).

Unfortunately, there are only few experimental claims in this debate. Most of these assumptions are based on thought experiments and cannot be experimentally studied. Actually, most of the authors still wonder whether epigenetic inheritance really plays a significant role in evolution on the long term. For instance, according to Pál and Hurst (2004): “*Although the role of epigenetic inheritance during development is generally accepted, it is much less clear whether heritable epigenetic variation can play a significant role during evolution*”.

To experimentally address this issue, we propose to use an *in silico* experimental evolution approach. By using an artificial life model, we are able to simulate evolution with and without epigenetic inheritance and to compare the outcomes. To this aim we used the RAevol model, developed in our team (Beslon et al., 2010a,b) to study the evolution of genetic regulation networks. We modified RAevol to allow for the vertical transmission of proteins from a parent to its offspring. In other words, in addition to transmitting its regulation network to its offspring (statically encoded on its genome), a parent also transmits the initial state of this network. Note that this represents only one of the various epigenetic inheritance mechanisms that exist in Nature (e.g. DNA methylation, chromatin structure, protein misfolding...). Furthermore, transmission of proteins is a weak epigenetic transmission because it is easily reversible in case of a plastic response. Nevertheless we show that this mechanism is strong enough to observe a significant change in the evolutionary dynamics.

This paper is organized as follows: In the next sections, we briefly present the RAevol platform and the way epigenetic inheritance is implemented. We then present our experimental design followed by the results of the simulations which are then discussed. Finally, we conclude and describe our future research directions.

The Aevol - RAevol platform

Aevol is an *in silico* experimental evolution (Hindré et al., 2012; Batut et al., 2013) or digital genetics (Adami, 2006) platform. It was designed to study how the evolutionary conditions shape the molecular structure of an evolving organism (e.g., DNA length, genes number, operonic structures...) due to direct and indirect selective pressures. In

Aevol, a population of individuals evolves through a classical mutation-selection process. The specificity of Aevol lies in the genotype-to-phenotype mapping which finely models what is observed in bacteria. A circular double-stranded DNA sequence is transcribed into a set of mRNAs. These mRNAs are then parsed in search for Coding DNA Sequences (CDSs – the “genes”) that are translated into proteins through an artificial genetic code. Finally, the proteins are combined to compute the individual’s phenotype. Since it is technically impossible to simulate a realistic cellular model, we use a simplified representation. In Aevol the phenotype of an organism is an $[0; 1] \rightarrow [0; 1]$ mathematical function that represents its metabolism. The fitness is then directly computed by comparing the phenotype with a target function representing the environment. In computational words, Aevol organisms must fulfill a curve-fitting task by combining kernel functions (the proteins) translated from their genomes.

Aevol has been extensively described elsewhere and we refer the reader to previously published work for a complete description of the model and the results obtained so far (Knibbe et al., 2007, 2008; Parsons et al., 2010; Batut et al., 2013; Misevic et al., 2015).

RAevol is an extension of Aevol (Beslon et al., 2010a,b). It uses the same genome model and the same genetic code. However, in RAevol, proteins are able to act as transcription factors (TFs) beside their metabolic activity. When acting as a TF, a protein may up- or down-regulate the transcription of other genes, ultimately controlling the concentration of the proteins encoded by these genes. In other words, in RAevol, each individual owns a genetic regulation network that may dynamically modify its phenotype depending on the environmental conditions. Importantly, in RAevol, the phenotype of an organism is no longer a static function (as it is in Aevol). Rather, it becomes a dynamic function that can be evaluated (by comparing it to a target function) at different time steps during what can now be considered as the “life” of the individual.

Technically, RAevol extends Aevol by adding a “transcriptional regulatory code”. In the model, the secondary structure of the proteins (*i.e.*, their sequence of Amino-Acids) may contain small motifs that are able to virtually bind on DNA sequences with an affinity that depends on the matching between the motif and the DNA sequence¹. If a protein is able to bind upstream or downstream from an mRNA promoter, then it respectively enhances or represses it. This results in an up- or down-regulation of the genes transcribed on this mRNA. Once the regulation network has been decoded from the genome, the transcription levels are

¹More precisely, the affinity is computed thanks to a regulatory matrix that associates 4AA motifs with 20 nucleotides-long sequences. This matrix is randomly drawn at the beginning of a simulation. Here the same matrix is used for all simulations.

used to parametrize a set of synthesis-degradation equations (one for each protein) and the dynamic of the network can be simulated step by step, enabling the computation of the dynamic phenotypic function during the life of the individual. In RAevol, the target phenotype may change during the life of the individual, either deterministically or randomly (Figure 1). The individual must then dynamically adapt to the current target by switching between different stable states of its regulation network. To this aim, evolution should optimize the “epigenetic landscape” encoded by the regulation network and create pathways between the different local minima associated to the target phenotypes.

All the details of the RAevol model can be found in (Beslon et al., 2010a,b). Classically, in RAevol, when an individual is created, all the proteins’ concentrations are initialized at their basal levels (*i.e.*, the level given by their promoter in the absence of regulation). For the needs of the present experiments, we have introduced a new option in the model: protein inheritance. When this option is selected, any new organism created in RAevol inherits its ancestor’s genome (possibly mutated) but it also inherits its ancestor’s proteins (*i.e.*, the Amino-Acid sequences and the corresponding concentration values). Conversely, all the proteins encoded in the organism’s genome are initialized at zero concentration. Then, due to the synthesis-degradation process, the inherited proteins are progressively degraded while the translated ones reach their steady state (see Figure 2, top-right panel). If a given CDS has been correctly transmitted from the parent to the offspring (*i.e.*, its coding sequence has not been altered by a mutation or a chromosomal rearrangement), this process results in a constant activity since the inherited protein is exactly replaced by the produced one. However, when a CDS undergoes a mutation, the offspring inherits a protein that it is not itself able to synthesize and produces a protein its parent was not able to produce. This may interfere with the regulation network convergence, hence altering the individual’s phenotype. As a consequence, in case of protein inheritance, the fitness of an organism is not only determined by its genome but also by its inherited cytoplasmic material. Both inheritance processes may thus interfere and we can study experimentally how this interference can impact evolution.

Experimental design

As argued in (Pál, 1998; Jablonka et al., 1995), the dynamic of the environment can be a crucial parameter for the contribution of epigenetic inheritance to evolution. Thus, in order to test whether epigenetic inheritance influences the evolutionary dynamics, we designed 2 series of scenarios. One will test the effect of protein inheritance in a variable environment and the other, in a constant environment.

In both scenarios the individuals live for 20 time steps (Figures 1 and 2). However, in the constant environment, the

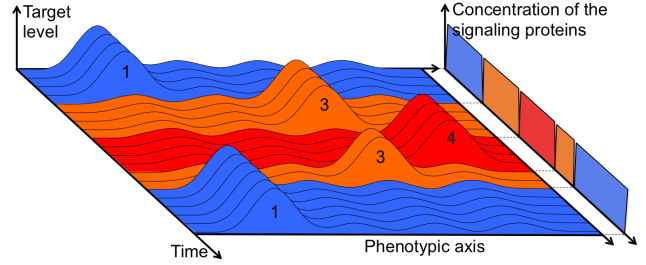


Figure 1: Example of a variable environment. The three environmental conditions are displayed respectively in blue, orange and red. They alternate randomly with a switching probability of 10% at each time step. Individuals life-duration is constant, fixed to 20 time steps. Each condition is also characterized by a signaling protein that can trigger a plastic response.

environmental conditions are fixed along the whole evolutionary process. More precisely, the phenotypic target is the sum of four Gaussians regularly spread along the phenotypic axis. On the opposite, in the case of a variable environment, we designed three different environmental conditions that alternate randomly with a switching probability of 10% at each time step (including the first one). In each of these conditions, one of the four Gaussians is active while the three others are inactive (Figure 1). In such an environment individuals can adapt by plasticity since the same conditions are encountered regularly (though randomly). Moreover, since the birth condition of an individual has a 0.9 probability to be its parental one, protein inheritance may reduce the lag time of the plasticity response. We will thus be able to test the theory proposed by Jablonka et al. (1995), to evaluate its contribution to evolution and to seek for additional mechanisms.

To summarize, we have four situations to simulate: variable environment with protein inheritance (V_i), variable environment without protein inheritance (V , control for V_i scenarios), constant environment with protein inheritance (C_i) and constant environment without protein inheritance (C , control for C_i scenarios).

For each scenario, we launched 4 independent simulations. In each simulation, a population of 1,000 individuals was let evolve for 300,000 generations under a mild mutation rate. All parameters were the same for all the simulations and were chosen according to previous experiments (Mutation rates: 5.10^{-6} mut/bp/generation for all kinds of point mutations. Rearrangement rates: 5.10^{-5} rearr/bp/generation for all kinds of rearrangements. Fitness proportionate selection. Selection coefficient: 750). Without epigenetic inheritance these parameters were known to lead to a medium complexity organisms with a good final fitness. In particular, they always lead to the emergence of

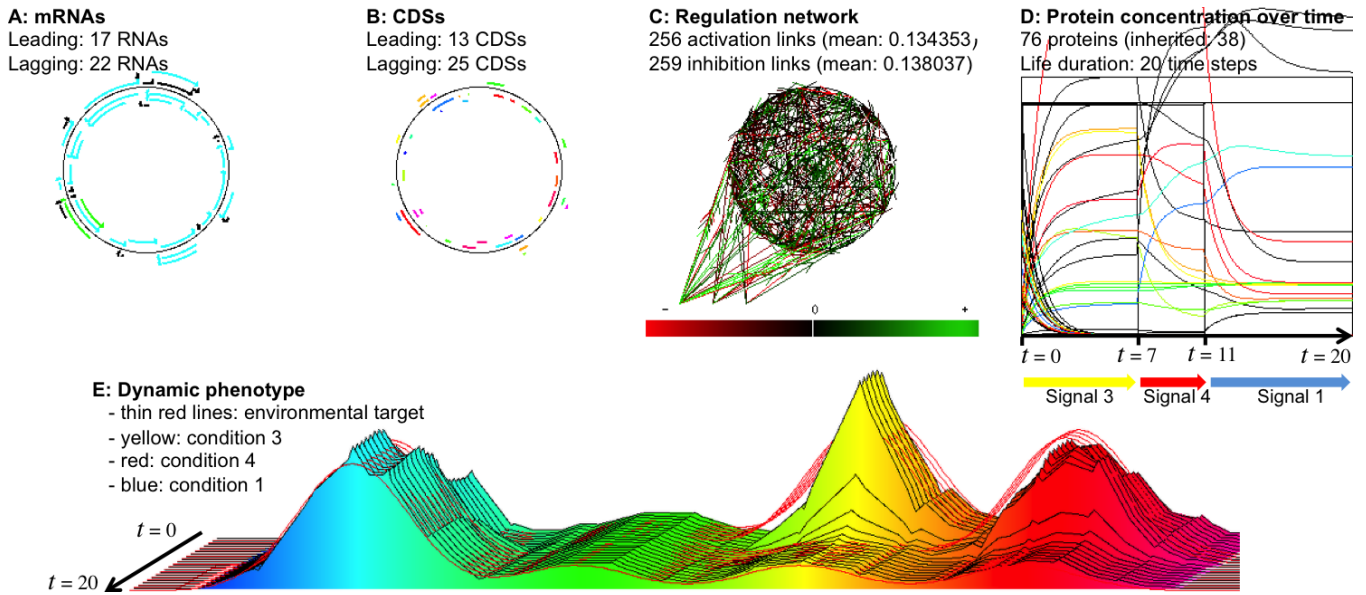


Figure 2: Example of an evolved RaeVol individual after 300,000 generations in V_i conditions (variable environment and epigenetic inheritance). **A**: genome and mRNAs (colors code for the basal transcription rate); **B**: genome and CDSs (colors indicate the function of the gene product); **C**: Regulation network (arrows indicate links between genes and mRNAs; colors indicate the weight of the link. Note the three external signals displayed at the bottom-left of the network); **D**: protein concentrations during the life of the individual (colors indicate the function of the protein). **E**: Dynamic phenotype of the individual (same color code as panels **B** and **D**; thin red lines indicate the target at each time step). This individual is born in condition 3. At time 7 the environment switches to condition 4 and at time 11 it switches to condition 1 (see Figure 1 for a detailed explanation). At each environmental change the regulation network modifies the transcription levels and thence protein concentrations, resulting in a phenotypic adaptation. Protein inheritance is clearly visible on panel **D**: 38 proteins are inherited (initialized at concentrations inherited from the ancestor but rapidly dropping to zero) and 38 are produced during the individual's life (initialized at concentration zero but rapidly reaching their steady-state).

a genetic regulation network as early as 50,000 generations. These parameters (and the relatively low number of repetitions) were chosen as a compromise between long enough evolution time and short enough computation time. Indeed, these simulations already last for more than 1 month on a 16-core 3GHz computer. It is important to mention that there is still room for directional evolution after 300,000 generations but previous experiments have shown that the main tendencies are already observed at this stage with the used parameter set. Indeed, during preliminary tests, we ran similar simulations for 800,000 generations without observing dramatic changes in the last period.

Results

All 16 simulations resulted in the emergence of organisms of medium complexity with most genomes ranging from 2,000 to 3,000 bp and approximately 30 genes (Figures 3.a and 3.b). Moreover, all the individuals were able to regulate their transcription and to efficiently respond to the environment changes (in the case of scenarios V/V_i). Figures 2 and 4 show the best final V_i individual and its genetic regulation network. Note the very fast proteome reorganization after

each environment variation (Figure 2, top-right panel).

As RAevol is a precise model of genome evolution, we are able to check whether protein inheritance has a major impact on the genome structure after 300,000 generations. As Figure 3 shows, there is no significant difference in neither the genome structure nor the genetic regulation network between scenarios with or without protein inheritance.

Conversely it is interesting to note that we do observe differences between scenarios C/C_i and V/V_i for the genome size (Figure 3.b) and average degree of the regulation network (Figure 3.c) but not for the number of genes (Figure 3.a). This strikingly contrasts with the classical hypothesis (see (Casjens, 1998) for example) that postulates that a link exists between the complexity of the environment and the size of the regulation network.

Here, the complexity of the environment is clearly higher for scenarios V/V_i than for scenarios C/C_i . However, the sizes of the regulation networks are similar (Figure 3.a) although the connectivity is lower for scenarios C/C_i (Figure 3.c). Moreover, the difference in molecular complexity observed here are orders of magnitude lower than what was previously observed in the same model when varying

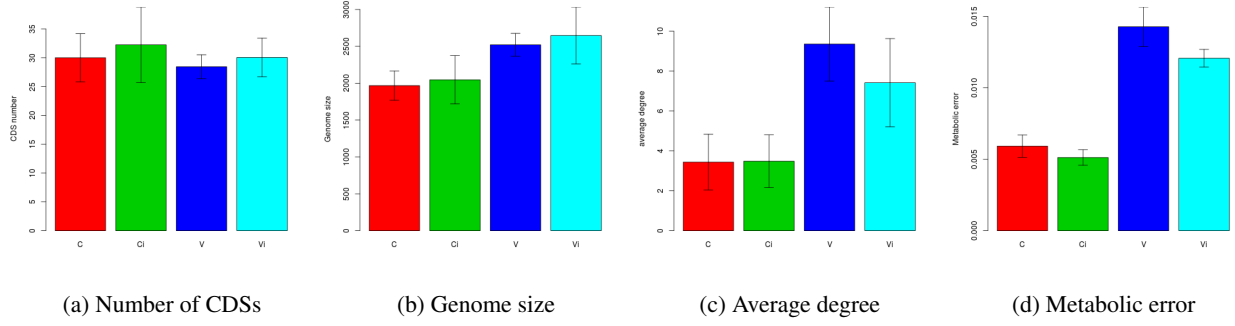


Figure 3: Comparison of the gene number, genome size, average degree of the regulation network and metabolic error of the best individual for the 4 scenarios at generation 300,000 (mean values of the 1000 last generations). Error bars show the standard deviation between the 4 simulations of a scenario. Red: scenario C , green: C_i , blue: V , cyan: V_i .

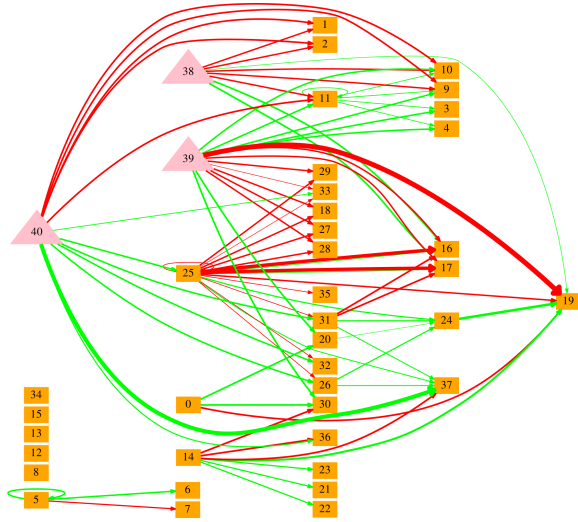


Figure 4: Genetic regulation network of the best individual in scenario V_i . Orange boxes represent genes; pink triangles represent the 3 signal proteins. Red (resp. green) arrows represent inhibition (resp. activation) links. Arrows weight represents the intensity of the link. For the sake of clarity, only links whose weight are at least 33% of that of the strongest link are represented.

the mutation rates rather than the environmental complexity (Beslon et al., 2010b). This observation raises very interesting questions regarding the origin of molecular complexity.

Although no clear molecular difference was observed between the protein inheritance scenarios (C_i/V_i) and their control (C/V), Figure 3.d shows a slight difference between them in terms of the individuals' performances. Indeed, the comparison of the mean metabolic error² of each scenario

²In Aevo/RAevo the adaptation of an individual is measured by its "metabolic error" *i.e.*, the total difference between its phenotypic function and the target function. Note that the lower the

shows us that scenarios V_i and C_i have lower metabolic error than scenarios V and C respectively (note that statistical accuracy cannot be reached with only 4 repetitions).

To understand this result, one can look at the evolution of metabolic error over time. Figure 5 clearly shows that protein inheritance accelerates the decrease of the metabolic error during the second phase of evolution (from 150,000 to 300,000 generations). However when looking qualitatively at the curves, we can notice that the evolutionary dynamic is clearly different with or without protein inheritance. In particular scenario C_i is characterized by long stasis phases separated by avalanches of favorable mutations, a situation that is not observed for scenario C . On the opposite, scenario V_i shows large fluctuations of metabolic error during the first phase of evolution (generation 0 to 50,000). These fluctuations initially slow-down evolution but scenario V_i ultimately catch up with scenario V because of a more regular evolution from generations 50,000 to 300,000. We can thus propose the hypothesis that the effects of protein inheritance are different depending on the evolutionary phase. Indeed, it seems to have a negative impact during the first evolution period (*i.e.*, when the network is being constructed by node recruitment) while its impact is clearly beneficial in the second period (*i.e.*, when the network structure is mostly stable but when the links are being optimized).

This hypothesis is confirmed when observing the evolution of the mean relative advantage given by protein inheritance over time. Indeed, Figure 6 confirms that the impact of protein inheritance is different in a constant environment and in a variable environment and that protein inheritance influence shows strong variations along the evolutionary process. For scenarios C/C_i , the advantage is slightly negative until generation 200,000 but with large fluctuations. Then it becomes stable and positive at almost 14%. This behavior is a direct consequence of the long evolutionary stasis phases

metabolic error, the better the individual (contrary to the usual fitness measures).

previously observed for scenario C_i (Figure 5). Moreover these phases are separated by large drops in the metabolic errors. Thus the relative advantage/disadvantage of protein inheritance is driven by stochastic events (the beneficial mutation avalanches). Given the low number of repetitions, this results in the large fluctuations observed on figure 6..

For scenarios V/V_i , we observe a huge disadvantage of protein inheritance that reaches -40% at generation 20,000. However this is rapidly compensated and protein inheritance becomes favorable as soon as generation 70,000. This is due to large fluctuations in the metabolic error in two repetitions of scenario V_i (Figure 5). Once these fluctuations stop, all the simulations rapidly reach a higher fitness than in scenario V . Interestingly, similar fluctuations are observed for the evolution of the number of CDSs (Figure 7). This confirms that the negative impact of protein inheritance is linked to the process of node recruitment by the regulation network.

As shown previously, during the second phase of the evolutionary process, protein inheritance seems to favor evolution whatever the conditions. However the advantage it provides can be due either to some facilitating process that would accelerate evolution or to a more “mechanical” effect: As explained above, protein inheritance can directly decrease the metabolic error because it enables transmission of an acquired phenotype. This is likely to be favorable when there is a high probability that an offspring faces the same conditions as its parents. In our simulations, this effect is likely to play an important role since the probability for an individual to begin its life facing the same environment its parent ended its life with is 90% in scenarios V/V_i and, obviously, 100% in scenarios C/C_i .

In order to evaluate the contribution of this effect, we measured the proportion of metabolic error that is due to the delay between an individual initialization and the moment when its regulation network reaches its steady state, for all the possible environmental conditions. At generation 300,000 we found a relative advantage of 7.5% (resp. 0.5%) for scenario V_i (resp C_i) over scenario V (resp C). Note that our estimation process overestimates the contribution of this effect since it does not take into account the possibility of an environmental switch immediately before or after the individual initialization (*i.e.*, the possibility that the environmental conditions change before the network reaches its steady state). As the mean relative metabolic error gain due to protein inheritance was 16,7% for scenarios V_i/V and 14,4% for scenarios C_i/C at generation 300,000, we can conclude that the final advantage observed for scenarios C_i and V_i is a combined effect of the direct transmission of an efficient acquired phenotype and of another yet to be discovered “facilitating” effect that accelerates fitness gain.

Discussion

We have presented two main results of our experiments. First, protein inheritance increases the long term fitness gain

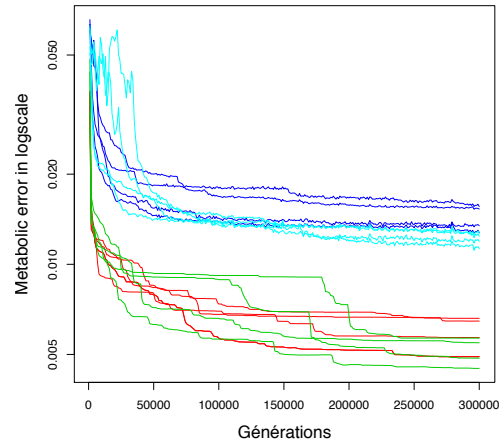


Figure 5: Evolution of the metabolic error of the best individual in scenarios C (red), C_i (green), V (blue) and V_i (cyan). For scenarios V and V_i the environmental condition changes randomly. This create random variation of the metabolic error between generations even for a same individual. To avoid this effect each point of the curve is the mean value of the 1,000 previous generations.

during evolution for scenarios C/C_i and V/V_i . Second, it has a negative effect during the first phase of the simulations: it generates instability in scenario V_i and clogging before fitness improvement for scenario C_i ³. In this section, we propose hypotheses to explain both observations.

Protein inheritance leads to major evolution impediment in the first phase of the simulations (*i.e.*, during the phase of gene acquisition and regulation network construction). This can be explained by a dynamic conflict between the two forms of inheritance that interact in the model. Indeed, the expression of a mutation does not follow the same timing for genetic inheritance and for protein inheritance. Indeed, if a non-silent mutation occurs at generation n , it leads to the emergence of a new phenotype at generation $n + 1$. However this new genotype will be expressed in the epigenetic context transmitted from generation n (*i.e.*, in an epigenetic context that does not include the mutation). Things are different at generation $n + 2$ where both the genetic inheritance and the epigenetic inheritance transmit the mutation. Depending on the genetic background and on the kind of mutation, this effect can be almost neutral but it can also have important consequences if the genetic and epigenetic contributions of the mutation are of opposed sign. A mutation with a positive genetic effect and a negative epigenetic effect could be beneficial at generation $n + 1$ (and thus be selected for) while being deleterious at generation $n + 2$. Similarly, a mutation

³The scenarios C and C_i can be compared by measuring the standard deviation of the derivative of the fitness. We find that the standard deviation is twice higher in scenario C_i than in scenario C (3.7710^{-7} vs. 1.6310^{-7}).

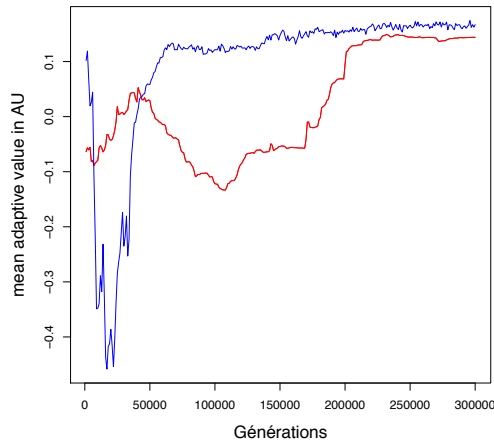


Figure 6: Evolution of the relative advantage of the protein inheritance in scenarios V/V_i (blue) and C/C_i (red). The relative advantage is computed as the difference of metabolic error divided by the mean metabolic error between the scenarios with and without protein inheritance.

with a negative genetic effect and a positive epigenetic effect could be counter-selected at generation $n + 1$ while it would have been favorable later.

This mechanism is likely to have a different impact on scenario V_i and C_i . Indeed, for scenario C_i , the target phenotype is constant and the genetic regulation network is not mandatory to survive (although it may be useful to finely tune the protein concentrations). Consequently, the interference between the two modes of selection adds noise to the evaluation function and may impede the fixation of favorable mutations but it is not severely detrimental. On the opposite, for scenario V_i , the regulation network is mandatory to survive in the three conditions that alternate in the environment. Moreover, evolution must create some pathways in the epigenetic landscape for the individuals to be able to switch from one target to another. There is thus a permanent competition between two evolutionary strategies: evolving a complex network that is able to efficiently switch between the environmental conditions, or keeping the population polymorphic (*i.e.*, to maintain subpopulations that only fit one of the environmental conditions). Given the switching probability, the regulation strategy is much more efficient but it is also much more affected by the interference between genetic inheritance and epigenetic inheritance since an efficient regulation network may be counter-selected “simply” because it is not initialized in the correct basin of attraction. Ultimately, this delays the moment when the regulation strategy invades the population, resulting in large fluctuations of the metabolic error.

Interestingly, this effect is also likely to depend on the kind of mutations. Indeed, if a mutation only affects an existing regulation link or an existing CDS, the consequences

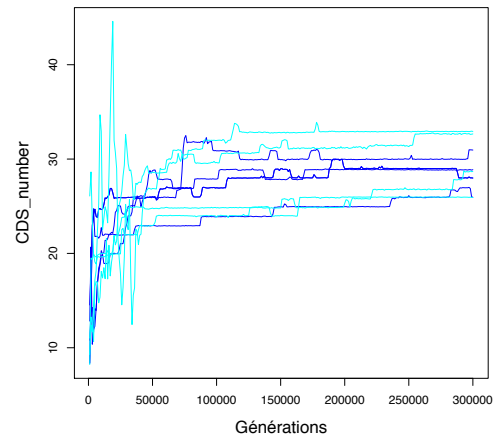


Figure 7: Evolution of the number of genes of the best individual in scenario V (blue) and V_i (cyan). As for figure 5, the curves have been smoothed.

on the epigenetic inheritance are likely to be mild (since there is only quantitative variations). On the opposite, if the mutation corresponds to the creation (respectively deletion) of a gene, this can have dramatic consequences since the mutation will add (respectively remove) a protein at generation $n + 1$. In such conditions, the interference between genetic and epigenetic inheritance is maximal. Indeed, a simple comparison between Figures 5 and 7 clearly shows that the fluctuations of metabolic error occur simultaneously with large fluctuations of the number of genes. This could explain why epigenetic inheritance is negative in the first evolutionary phase (when many genes are acquired in a “few” generations) but positive in the second phase.

The positive effect of epigenetic inheritance on the long term is much more straightforward. First, as we have seen above, the advantage of epigenetic inheritance must be decomposed between a direct advantage (the transmission of the epigenetic memory reduces the lag phase when the environmental conditions of the parent and of the offspring are correlated – which is often the case in our simulation) and an indirect advantage that accelerates evolution. This indirect advantage is easily explained in terms of pathways in the epigenetic landscape. Indeed, in scenario V_i , the epigenetic landscape must contain 3 pathways to enable switching between the different environmental conditions: $1 \leftrightarrow 2$, $1 \leftrightarrow 3$ and $2 \leftrightarrow 3$ (each of them triggered by a signaling protein). However, in scenario V , the epigenetic landscape must contain 6 pathways: $1 \leftrightarrow 2$, $1 \leftrightarrow 3$, $2 \leftrightarrow 3$ plus $I \rightarrow 1$, $I \rightarrow 2$ and $I \rightarrow 3$ (I being the initial state of the regulation network). Consequently, the genetic network is much more constrained for the individuals in scenario V than for those in scenario V_i . The situation is similar for individuals in scenario C that must evolve an epigenetic pathway from their initial state to their steady state while in scenario C_i ,

individuals are always initialized at the steady state.

Conclusion and future work

In this paper, we experimentally addressed the question of the impact of non-genetic inheritance on evolution (*i.e.*, whether it could accelerate or decrease the rate of evolution). To evaluate this question, we used an *in-silico* experimental evolution platform, RAevol, to test 4 scenarios: 2 with a static environment (C and C_i) and 2 with a dynamic one (V and V_i). For each, we launch 2 types of simulations: 2 with protein inheritance (C_i and V_i) and 2 without (C and V). Our simulations show that protein inheritance indeed increases the rate of evolution on the long term but that it can have a strong negative effect at the beginning of the evolutionary process by generating instability in V_i scenario and clogging before fitness jumps for scenario C_i . Finally, we discussed these results and proposed that the evolutionary consequences of protein inheritance are a complex composition of three mechanisms: (*i.*) interference between genetic and protein inheritance due to their different transmission delays. This creates noise in the selection process and may delay the emergence of an efficient regulation network; (*ii.*) epigenetic memory that reduces the lag time when parents and offspring environments are correlated and (*iii.*) simplification of the epigenetic landscape.

The former effect was shown to be strongly deleterious in our simulations but only at the beginning of the evolution. It is then rapidly compensated by the two latter, resulting in a long-term benefit of protein inheritance on the evolution of our artificial organisms.

To the best of our knowledge, this work is the first attempt to explore experimentally the question of the impact of epigenetic inheritance on evolutionary dynamic. We now need to validate statistically our results by running more experimentation and to test other environmental conditions to confirm our mechanistic hypothesis (*e.g.*, by letting individuals evolve in an environment where variation is correlated with the reproduction).

Availability

Aevol is available under GPL licencing at the project website: <http://www.aevol.fr>. RAevol is currently in beta-version and is available upon request from the authors.

Acknowledgements

We would like to warmly thank David P. Parsons for careful reading of the manuscript. This research has been supported by EU-FET grant EvoEvo (ICT-610427).

References

- Adam, M., Murali, B., Glenn, N., and Potter, S. (2008). Epigenetic inheritance based evolution of antibiotic resistance in bacteria. *BMC Evolutionary Biology*, 8:12.
- Adami, C. (2006). Digital genetics: unravelling the genetic basis of evolution. *Nature Reviews Genetics*, 7(2):109–118.
- Batut, B., Parsons, D. P., Fischer, S., Beslon, G., and Knibbe, C. (2013). In silico experimental evolution: a tool to test evolutionary scenarios. *BMC Bioinformatics*, 14 (S15):S11.
- Beslon, G., Parsons, D. P., Peña, J.-M., Rigotti, C., and Sanchez-Dehesa, Y. (2010a). From digital genetics to knowledge discovery: Perspectives in genetic network understanding. *Intelligent Data Analysis Journal*, 14(2):173–191.
- Beslon, G., Parsons, D. P., Sanchez-Dehesa, Y., Peña, J.-M., and Knibbe, C. (2010b). Scaling laws in bacterial genomes: A side-effect of selection of mutational robustness. *BioSystems*, 102(1):32–40.
- Bonduriansky, R. and Day, T. (2009). Nongenetic inheritance and its evolutionary implications. *Annual Review of Ecology, Evolution, and Systematics*, 40:103–125.
- Casjens, S. (1998). The diverse and dynamic structure of bacterial genomes. *Annual review of genetics*, 32(1):339–377.
- Danchin, E., Charmantier, A., Champagne, F. A., Mesoudi, A., Pujol, B., and Blanchet, S. (2011). Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews Genetics*, 12(7):475–486.
- Hindré, T., Knibbe, C., Beslon, G., and Schneider, D. (2012). New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nature Reviews Microbiology*, 10:352–365.
- Jablonka, E., Oborny, B., Molnar, I., Kisdi, E., Hofbauer, J., and Czarán, T. (1995). The adaptive advantage of phenotypic memory in changing environments. *Philosophical Transactions of the Royal Society B*, 350:133–141.
- Knibbe, C., Coulon, A., Mazet, O., Fayard, J.-M., and Beslon, G. (2007). A long-term evolutionary pressure on the amount of noncoding DNA. *Molecular Biology and Evolution*, 24(10):2344–2353.
- Knibbe, C., Fayard, J.-M., and Beslon, G. (2008). The topology of the protein network influences the dynamics of gene order: From systems biology to a systemic understanding of evolution. *Artificial Life*, 14(1):149–156.
- Lachmann, M. and Jablonka, E. (1996). The inheritance of phenotypes: an adaptation to fluctuating environments. *Journal of Theoretical Biology*, 181:1–9.
- Misevic, D., Frénay, A., Lindner, A. B., and Taddei, F. (2015). Shape matters: Lifecycle of cooperative patches promotes cooperation in bulky populations. *Evolution*, 69(3):788–802.
- Pál, C. (1998). Plasticity, memory and the adaptive landscape of the genotype. *Proceedings of the Royal Society of London B: Biological Sciences*, 265:1319–1323.
- Pál, C. and Hurst, L. D. (2004). Epigenetic inheritance and evolutionary adaptation. In *Organelles, Genomes and Eukaryote Phylogeny: An Evolutionary Synthesis in the Age of Genomics*. CRCpress.
- Parsons, D. P., Knibbe, C., and Beslon, G. (2010). Importance of the rearrangement rates on the organization of transcription. In *Proceedings of Artificial Life XII*, pages 479–486.